

NOVA COLLEGE-WIDE COURSE CONTENT SUMMARY ITD 240 – Machine Learning II (3 CR.)

Course Description

Examines theory, algorithms, applications, and issues within the subfield of pattern recognition and machine learning, including feature engineering and extraction, supervised and unsupervised learning. Focuses on theory and practice, with coverage of underlying mathematical and heuristic concepts. Part II of II. Lecture 3 hours per week.

General Course Purpose

The purpose of this course is to give the student competence in describing, choosing, training, testing, and evaluating the efficacy and applicability of various machine learning algorithms and methods, using case studies and real-world applications.

Course Prerequisites/Corequisites

Prerequisite: ITD 140 Machine Learning I or division approval

Recommended: ITP 150 Python I or intermediate programming experience

Course Objectives

Upon completing the course, the student will be able to:

- Describe and identify basic artificial intelligence approaches.
- Describe and differentiate between supervised and unsupervised learning techniques.
- Identify and explain basic types of machine learning algorithms for both supervised and unsupervised machine learning.
- Describe and apply feature extraction and engineering techniques.
- Explain scenarios where supervised or unsupervised learning would be appropriate.
- Apply supervised and unsupervised algorithms to problems in the form of case studies.
- Demonstrate basic proficiency in data preparation and machine learning training and testing by both using GUI applications and tools as well as writing code in Python.
- Explain and apply feature extraction and engineering techniques.
- Explain and demonstrate basic proficiency in evaluating and tuning models.

Major Topics to be Included

- Artificial intelligence; overview, history
- Machine learning generally
- Supervised learning – regression and classification
- Unsupervised learning – clustering, dimensionality reduction
- Feature engineering and extraction
- Optimization and loss functions
- Decision trees
- Neural networks
- Noise, overfitting, bias
- Model validation and evaluation metrics
- Model and hyperparameter tuning

Student Learning Outcome

- A. Identify artificial intelligence terminology
 1. Define and explain the purpose of Artificial Intelligence (AI); define AI Winter

2. Define and explain the purpose of Artificial Narrow Intelligence (ANI)
3. Define and explain the purpose of Artificial General Intelligence (AGI)
4. Explain the capabilities of Artificial General Intelligence (AGI)
- B. Identify, define and explain the purpose of machine learning
- C. Identify feature engineering terminology
 1. Define regression imputation
 2. Define multiple imputation
 3. Define nominal attributes
 4. Define ordinal attributes
- D. Define supervised learning terminology
 1. Define and explain the purpose of supervised learning
 2. Demonstrate ability to identify supervised learning algorithms and identify appropriate applications
 3. Define regression as a supervised learning prediction task
 4. Demonstrate the ability to identify when it is appropriate to use regression
 5. Define and explain the purpose of decision tree learning
 6. Define classification as a supervised learning prediction task
 7. Demonstrate the ability to identify appropriate applications of classification
 8. Define k-nearest neighbors method for classification prediction tasks
 9. Define bias within linear models
 10. Define least squares for linear regression modeling
 11. Define entropy
 12. Define and explain the purpose of the Stochastic Gradient Descent model
 13. Define and explain the purpose of deep learning
 14. Define and explain the purpose of support vector machines (SVMs)
 15. Define and explain the purpose of a Bayesian Network
 16. Define bias, variance, overfitting, underfitting
 17. Define bias-variance tradeoff
 18. Define and explain the purpose of hyperparameters
 19. Demonstrate the ability to explain the impact of hyperparameters on a model
 20. Understand the impact of hyperparameters on complexity
 21. Define and explain the purpose of weights
 22. Apply supervised learning to analyze and solve real world problems through case studies
- E. Define unsupervised learning terminology
 1. Define unsupervised learning
 2. Define clustering
 3. Define and explain the purpose of anomaly detection
 4. Define and explain the purpose of dimensionality reduction
 5. Demonstrate the ability to identify when dimensionality reduction is appropriate
 6. Define and explain the purpose of k-means clustering
 - a. Define inter-cluster distance
 - b. Define centroids
 7. Define and explain the purpose of tuning
 8. Apply unsupervised learning to analyze and solve real world problems through case studies
- F. Apply ML theory and applications to analyze and solve real world problems through case studies

Required Time Allocation per Topic

To standardize the core topics of this course, the following student contact hours per topic are required. Each syllabus should be created to adhere as closely as possible to these allocations. Topics are not necessarily to be taught in the order shown.

There are normally 45 student contact-hours per semester for a three-credit course (14 weeks of instruction, excluding final exam week: $14 \times 3.2 = 45$ hours). Sections of the course offered in alternative formats (i.e. not standard 15-week) still meet for the same number of contact hours. The final exam is not included in the time table.

The quickly evolving nature of artificial intelligence—and machine learning in particular—means that some content noted in this document may be superseded or made obsolete. As such, it is important to include such changes in individual syllabi. Additionally, time is allocated for additional and optional topics to provide flexibility to instructors in tailoring the course to special needs or resources.

Topics	Hours	Percentage
Artificial intelligence; overview, history; Machine learning generally	1	2%
Supervised learning – regression and classification	9	20%
Unsupervised learning – clustering, dimensionality reduction	9	20%
Feature engineering and extraction	10	23%
Optimization and loss functions	3	7%
Decision trees; Neural networks	3	7%
Noise, overfitting, bias; Model validation and evaluation metrics	3	7%
Model and hyperparameter tuning	2	4%
Testing to include quizzes, tests and exams (excluding final exam)	3	7%
Other Optional Topics	2	4%
Total	45	100%